DOCUMENT RESUME

ED 267 126	TM 860 219

AUTHOR	Chastain, Robert L.; Willson, Victor L.
TITLE	Estimation of Variance Components Using Computer Packages.
PUB DATE	[86]
NOTE	15p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Analysis of Variance; *Computer Software; Correlation; *Estimation (Mathematics); *Generalizability Theory; *Maximum Likelihood Statistics; Programing; Research Design; Statistical Data
IDENTIFIERS	*Repeated Measures Design; *Statistical Packages; Variance (Statistical); Wechsler Adult Intelligence Scale (Revised)

ABSTRACT

Generalizability theory is based upon analysis of variance (ANOVA) and requires estimation of variance components for the ANOVA design under consideration in order to compute either G (Generalizability) or D (Decision) coefficients. Estimation of variance components has a number of alternative methods available using SAS, BMDP, and ad hoc procedures. This study examines the effect of these methods on coefficeints for a large sample of subjects under complex designs. The Wechsler Adult Intelligence Scale-Revised national standardization sample provided the data base. Subject (n=1880) variables (age, gender, etc.) provided variety for the construction of complex factoral designs. Three designs were used to investigate variance component estimation and resultant g-coefficeints. The designs included: (1) a two-factor subject x item design; (2) a three-factor subject x item x age design; and (3) a four-factor subject x item x age x gender design. These designs allowed estimation of both within and between subject effects. Suggestions are made concerning equal and unequal situations. (Author/PN)

Estimation of Variance Components Using Computer Packages

Robert L. Chastain

Stanford University

Victor L. Willson

Texas A&M University

Running head: VARIANCE COMPONENTS

Abstract

Generalizability theory is based upon analysis of variance (ANOVA) and requires estimation of variance components for the ANOVA design under consideration in order to c ite either G (Generalizability) or D (Decision) coefficients (Cardinet, Tourneur, & Allal, 1976). Estimation of variance components has a number of alternative methods available using SAS, BMDP, and ad hoc procedures. This study examines the effect of these methods on coefficients for a large sample of subjects under complex designs. Suggestions are made concerning equal and unequal situations.

ESTIMATION OF VARIANCE COMPONENTS USING COMPUTER PACKAGES

Introduction

The theory of generalizability of behavioral and mental measurements depends upon the estimation of variances for various facets in the design under consideration. For example, the simplest generalizability (reliability) coefficient or g-coefficient based upon a subject-by-item design is estimated using the intraclass correlation. This statistic estimates the ratio of subject variance to subject plus error variance, but it is not the only statistic that can be computed to estimate that ratio. Several statisticians have derived alternate methods for variance components estimation, including Hemmerle and Hartley (1973) for maximum likelihood; Hartley, Rao, and LaMotte (1978) for the MIVQUE0 method; Gaylor, Lucas, and Anderson (1970) for the adjusted sum of squares method, and various quasi-F procedures of an ad hoc nature. Consequently, it is not known presently how estimated g-statistics vary in magnitude by use of various estimation procedures. Recently, Bell (1985) examined several procedures within SAS and concluded that MIVQUE0 was the most useful and efficient.

Although other statistical packages were not considered, nor were components calculated from unbalanced data, software problems concerning the estimation of variance components were brought to light. This paper answers some of the questions raised by Bell and suggestions are given concerning unequal n situations.

Theoretical Framework. The theory of generalizability due to Cronbach, Gleser, Nanda, and Rajaratnam (1972) was used as the definitive work for all designs employed. Intraclass correlation coefficients were calculated from the variance components based upon developments by Cardinet, Tourneur, and Allal (1976). The resulting g-coefficients were computed according to Rentz's (1980) rules for calculation.

METHOD

The Wechsler Adult Intelligence Scale-Revised national standardization sample provided the data base (Wechsler, 1981). It consists of 1880 subjects stratified by age, gender, race, occupation, education, urban-rural residence, and region of the U.S. In addition, several other variables were available, such as birthplace, handedness, and birth order. These variables provide great variety for the construction of

complex factoral designs. Three designs were used to investigate variance component estimation and resultant g-coefficients. The three designs were : (a) a two-factor subject x item design, (b) a three-factor subject x item x age design, and (c) a four-factor subject x item x age x gender design. These designs allowed estimation of both within and between subect effects.

Variance components and the resultant g-coefficients for the intraclass correlations were computed and calculated using the BMDP program 8V and SAS. BMDP8V uses the Cornfield and Tukey (1956) formulas to estimate variance components. The new SAS Version 5 procedure VARCOMP has four estimation methods: (a) TYPE1, (b) MIVQUE0, (c) ML, and (d) REML (SAS, 1985b). The TYPE1 method computes the Type 1 sum of squares for each effect (factor) and then solves the resulting system of equations (Gaylor, Lucas, & Anderson, 1970). The MIVQUE0 method uses a technique similar to the TYPE1 method except that the r[illegible]m effects are adjusted only for the fixed effects. This method affords a considerable timing advantage over the TYPE1 method (SAS, 1985b). The MIVQUE0 method is based on the Hartley, Rao, and LaMotte (197[illegible]) technique designed to produce estimates that are locally best quadratic unbiased estimates. The ML (maximum-likelihood)

method computes maximum-likelihood estimates of the variance components by using initial MIVQUE0 estimates and the W-transformation developed by Hemmerle and Hartley (1973). This procedure then iterates until the log-likelihood objective is satisfied or coverges. The REML (restricted maximium-likelihood) method is similar to the ML method except that the likelihood is separated into two parts: fixed effects and random effects (Patterson & Thompson, 1971). Initial MIVQUE0 estimates are iterated until convergence is met for the random effects part only.

Five random samples were selected from the WAIS-R data base (N=1880) to examine the stability of the estimates and to have both balanced and unbalanced designs. The WAIS-R standardization sample was randomly divided into four subsamples (N=470 each). These four subsamples allowed both equal and unequal cell sizes according to the number of factors investigated to examine the effects of unequal cell sizes on estimates of the variance components, g-coefficients, computation time, and memory required for computation. A fifth subsample (N=360) was created by randomly sampling so that there would be equal cell sizes for the two-, three-, and four-factor designs.

RESULTS

Although most results were obtained for the two-factor solutions, no ML or REML solutions were obtained for the three- and four-factor designs. The time required for ML and REML solutions under SAS for both equal and unequal sample size designs became prohibitively expensive (over 6 minutes executed c.p.u. and over 15 minutes calculated c.p.u.). Bell (1985) also found the VARCOMP procedure (methods of TYPE1 and ML) and the GLM procedure within SAS to be prohibited by requiring large (over 5 minutes c.p.u. time) amounts of time and memory which are expensive.

It was hoped by the present authors that the new SAS Version 5 might reduce the time and memory requirements using the new REML method in the VARCOMP procedure or the new REPEATED statement in the GLM procedure. The GLM REPEATED statement was investigated using the two-factor equal cell size subsample (N=360) and it required over 3 minutes c.p.u. and over 2000K of memory. The REML method also used large amounts of time and memory.

The time (in c.p.u.) and memory requirements for the different methods and procedures are shown in Table 1. For the MIVQUE0 and TYPE1 methods, more time and memory was needed as the number of factors increased.

Also, less time was needed for the equal cell size sample (EQSAM) than for the other four samples for each design (two-, three-, and four-factors). This should be true regardless of balanced versus unbalanced designs because there were less subjects in EQSAM compared to the other four subsamples.

The BMDP program 8V provided the most efficient estimates with balanced data. Interestingly, the BMDP8V program was even more efficient when accessed through SAS (SAS, 1985a). This finding was important because the data must be sorted carefully for data entry to be used in PMDP8V according to slowest and fastest moving indexes or levels of factors (Dixon, 1985). The data can be sorted easily with the SORT procedure prior to the BMDP procedure in SAS. Also, this finding allows for the analysis of SAS structured data sets.

Table 2 shows the estimated g-coefficients for the various procedures and methods, samples, and designs. For the two-factor solutions there was no variation to the third decimal place across procedures for EQSAM. The estimates varied from .908 to .931 across all five random samples. Since reliability estimates are often reported to two decimal places, the .91 to .93 difference is negligible.

The three-factor solutions provided estimates that

were very close to the two-factor solutions. These estimates varied from g-coefficients of .907 to .935. Although these estimates varied a little more than the two-factor solutions, the estimates were very similar across procedures and samples.

The four-factor solutions revealed more variation across procedures and samples than either the two- or three-factor solutions. This was due in large part to the MIVQUE0 method consistently estimating larger error variances. The procedures varied from .920 to .926 for EQSAM, but varied from .890 to .937 across all five random samples. Even these variations may not represent substantial differences in terms of efficiency or precision.

CONCLUSIONS

The major results of this study supplement Bell's (1985) results. For typical, large test data sets it will be extremely expensive to estimate g-coefficients if TYPE 1, ML, or REML methods of SAS's VARCOMP procedure are used to compute variance components. BMDP, while restricted to equal cell size designs, was efficient in its computations and required fractions of the time and memory used by those three methods. For many situations unbalanced designs will occur, however, with excessive loss of information accompanying

balancing. This situation can be assessed by using the MIVQUE0 method of SAS's VARCOMP procedure for the unequal cell size sample and then comparing these estimates with BMDP8V estimates of the reduced, balanced, equal cell size sample. It is unlikely that these two methods will give drastically different estimates for simple designs (two- or three-factors), but different estimates could result for designs with many factors and many levels.

It was anticipated that the new SAS Version 5 would help meet the need for efficient computation of within and between subject effects with the REPEATED statement of the GLM procedure. This has lowered the cost for repeated measures analysis, but is not efficient as applied to the calculation of variance components from Type 1 expected mean squares for generalizability theory.

References

Bell, J. F. (1985). Generalizability theory: The software problem. Journal of Educational Statistics, 10, 19-29.

Cardinet, J., Tourneur, Y., & Allal, L. (1976). The symmetry of generalizability theory: Applications to educational measurement. Journal of Educational Measurement, 13, 119-135.

Cornfield, J., & Tukey, J. W. (1956). Average values of mean squares in factorials. Annals of Mathematical Statistics, 27, 907-949.

Cronbach, L., Gleser, G., Nanda, M., & Rajaratram, N. (1972). The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: John Wiley and Sons.

Dixon, W. J. (1985). BMDP statistical software: 1985 printing. Berkely, CA : University of California Press.

Gaylor D. W., Lucas, M. L., & Anderson, R.L. (1976). Calculation of the expected mean squares by the abbreviated Doolittle and square root methods. Biometrics, 26, 641-656.

Hartley, H. O., Rao, J. N. K., & LaMotte, L. R. (1978). A simpler synthesis-based method of variance component estimation. Biometrics, 34, 233-242.

Hemmerle, W. J., & Hartley, H. O. (1973). Computing maximum likelihood estimates for the mixed AOV model using the W-tranformation. Technometrics, 15, 819-831.

Patterson, H. D , & Thompson, R. (1971). Recovery of inter-block information whwn block sizes are unequal. Biometrics, 15, 819-831.

Rentz, R. R. (1980). Rules of thumb for estimating reliability coefficients using generalizability theory. Educational and Psychological Measurement, 40, 575-592.

SAS Institute. (1985a). SAS user's guide: Basics. Cary, N. C.: Author.

SAS Institute. (1985b). SAS user's guide: Statistics. Cary, N. C.: Author.

Wechsler, D. (1981). The Wechsler Adult Intelligence Scale-Revised. New York: The Psychological Corporation.

Table 1

C.P.U. and Memory Requirements

Method	Factors	Average C.P.U.	Average Region
MIVQUE0	2	2.1 (1.4)*	620 K
	3	3.0 (1.9)	1120 K
	4	3.7 (3.3)	2000 K
TYPE1	2	33.6 (15.6)	620 K
	3	56.6 (34.1)	1270 K
	4	91.3 (72.7)	2050 K
REML	2	over 360	3000 K
	3	over 360	3000 K
	4	over 360	3000 K
ML	2	over 360	3000 K
	3	over 360	3000 K
	4	over 360	3000 K
BMDP	2	1.2	600 K
	3	1.5	600 K
	4	1.9	600 K
SASBMDP	2	0.4	600 K
	3	0.8	600 K
	4	0.8	600 K

* NOTE: Equal cell size sample (EQSAM) requirements

Table 2

G-coefficient Estimates

Two-factor Solutions

Procedure	EQSAM	One	Two	Three	Four
BMDP	.927	----	----	----	----
SASBMDP	.927	----	----	----	----
MIVQUE0	.927	.908	.914	.927	.931
TYPE1	.927	.908	.914	.927	.931
ML	.927	----	----	----	----
REML	.927	----	----	----	----

Three-factor Solutions

Procedure	EQSAM	One	Two	Three	Four
BMDP	.927	----	----	----	----
SASBMDP	.927	----	----	----	----
MIVQUE0	.927	.907	.915	.928	.935
TYPE1	.927	.908	.915	.928	.935

Four-factor Solutions

Procedure	EQSAM	One	Two	Three	Four
BMDP	.926	----	----	----	----
SASBMDF	.926	----	----	----	----
MIVQUE0	.920	.906	.890	.925	.934
TYPE1	.926	.907	.912	.926	.934